$$H(X|Y) \leq H(X)$$

$$H(X,Y) = H(X) + H(Y|X)$$

$$H(X,Y) \leq H(X) + H(Y)$$

$H(X) \leq \log_2 k$ if X takes on k distinct values

Proof:

$$H(X) = \sum_a P[X=a] \log_2\left(\frac{1}{P[X=a]}\right)$$

$$\leq \log_2\left(\sum_a \cancel{P[X=a]} \frac{1}{\cancel{P[X=a]}}\right) = \log_2(k)$$

Unicity Distance for Caesar

Assume that we have just intercepted N letters of ciphertext that was encrypted using a Caesar shift. How large does N have to be (on average) in order the uniquely determine the shift? Assume that the entropy of the english language is 3.2 bits.

We begin with the identity:

$H(C) \leq \log_2 26 \simeq$ 4.7

$$\boxed{H(K) = H(C) - H(M)}$$

$1.78 < H(M) < 4.7$

Assuming that each of the 26 keys is equally likely, we have

$$\boxed{H(K) = \log_2 26 \approx 4.7}$$

Assuming that each of the 26^N ciphertexts is equally likely, we have

$$\boxed{H(C) = \log_2 26^N = N \log_2 26 \approx 4.7N}$$

Therefore,

$H(K)$ $H(C)$ $H(M)$

$$\boxed{4.7 = 4.7N - 3.2N \Rightarrow N \approx 3.13}$$

Monoalphabetic Substitution

Assume that we have intercepted N letters of a ciphertext message that was encoded using a Monoalphabetic substitution and that the entropy of english is 2 bits. ← I have a larger N for monoalphabetic
per letter

Length of text	5	10	15	20	30	40	50
# of distinct letters	4	8	11	12	14	16	18

For instance, a typical english sample of 30 letters contains about 14 di erent letters. Thus the key for a Monoalphabetic substitution only permutes 14 letters. Therefore the number of keys is

guess $N \approx 30$

$$26 \times 25 \times \cdots \times 13 = (26-14+1)$$

$$(26-1+1)(26-2+1)(26-3+1)\cdots(26-14+1)$$

and not 26!.

Assuming that each key is equally likely, we have

$$\boxed{H(K) = \log_2(26 \times 25 \times \cdots \times 13) \approx 59.54}$$

Assuming that each of the 26^N ciphertexts is equally likely, we have

$$\boxed{H(C) = \log_2 26^N = N \log_2 26 \approx 4.7N}$$

Therefore,

$H(M) = 2N$

$$\boxed{59.54 = 4.7N - 2N \Rightarrow N \approx 22.05}$$

With $N \approx 22$ then the # distinct letters is not 14 but closer to 12 OR 13

Maybe # of keys =

$$26 \cdot 25 \cdot 24 \cdot \cdots \cdot 15$$

OR

$$26 \cdot 25 \cdot 24 \cdot \cdots \cdot 14$$

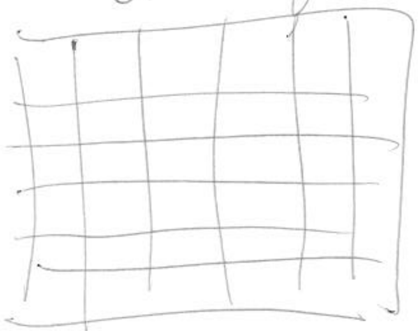
Try recalculating N for "this" estimation of $H(K)$

$$H(K) = \log_2(26 \cdot 25 \cdots \cdot 15)$$

ADFGVX - encipherment scheme

Key 1 filling of 6x6 square

Key 2 permutation of 18



$$H(K) = \log_2\left(18! \cdot \overset{=(36-1+1)}{36} \cdot 35 \cdot 34 \cdot \ldots \cdot \overset{=(36-26+1)}{11}\right)$$

$$= \log_2(18!) + \log_2\left(\frac{36!}{10!}\right)$$

$$= 52.5 + 116.3$$

$$H(C) = \log_2 6^N = N(1 + \log_2 3)$$

$$H(M) = 1.2\left(\frac{N}{2}\right)$$

$$N = \frac{52.5 + 116.3}{(1 + \log_2 3) - 1.2/2}$$